

REMARKS

The present application was filed on December 14, 1999 with claims 1-8. Claims 9-19 were added in an Amendment dated March 7, 2001. In the outstanding Office Action dated November 7, 2001, the Examiner has: (i) rejected claims 1-5, 9-12 and 18 under 35 U.S.C. §112, second paragraph as being indefinite; (ii) rejected claims 6 and 19 under 35 U.S.C. §102(b) as being anticipated by U.S. Patent No. 5,355,433 to Yasuda et al. (hereinafter "Yasuda"); (iii) rejected claims 7, 8, 16 and 17 under 35 U.S.C. §103(a) as being unpatentable over Yasuda; (iv) rejected claim 13 under §103(a) as being unpatentable over Yasuda in view of U.S. Patent No. 6,233,559 to Balakrishnan (hereinafter "Balakrishnan"); (v) rejected claims 6-8, 13, 16, 17 and 19 under §103(a) as being unpatentable over U.S. Patent No. 6,282,508 to Kimura et al. (hereinafter "Kimura"); (vi) indicated that claims 1-5, 9-12 and 18 would be allowable if rewritten to overcome the §112 rejection; and (vii) indicated that claims 14 and 15 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form. Applicants wish to thank the Examiner for his indication of allowable subject matter.

In this response, Applicants have amended claims 1 and 18 in a manner which Applicants believe addresses the Examiner's §112 rejection. Furthermore, Applicants traverse the §102 and §103 rejections for at least the reasons set forth below. Applicants respectfully request reconsideration of the present application in view of the above amendments and the following remarks.

As stated in Applicants' prior response dated August 27, 2001, reiterated herein for convenience, the present invention provides techniques "for contingent transfer and execution of spoken language interfaces" (present specification; page 2, lines 24-25). In a portable speech assistant (PSA), "a spoken language interface is defined in sets of user interface files. These are referred to as vocabularies files, prompt files, profiles and scripts depending on the role they play in structuring the interface" (present specification; page 3, lines 3-5). As used by the present invention, the term "spoken language interface" is intended to refer to the general act of speaking to a machine, listening to a machine, and/or interacting with a machine through utterances or audible expressions, and does not refer to a particular lingual type (e.g., English or Spanish).

An important aspect of the present invention is its ability to dynamically instantiate a new application and its spoken language interface (present specification; page 42, lines 5-7). It should be appreciated that the spoken language interface is a collection of operable features that allows a user to interact with the application. For example, user utterances may operate the features of the application, e.g., by supplying a reference to one or more events to be processed by the target application (present specification; page 3, lines 5-6). The term “event” is used by the present invention in a conventional sense in the context of event handling programs, e.g., event handling may be a feature of the application. These operable features, which are built into an application and are controlled at least in part by user utterances, are to be distinguished from data on which the application program acts.

Claims 6 and 19 have been rejected under §102(b) as being anticipated by Yasuda. Specifically, the Examiner contends that Yasuda teaches all of the elements of the invention, as set forth in the above claims. Applicants respectfully disagree with this contention. Yasuda is directed to a standard pattern comparing system “which recognizes data by comparing it with a standard pattern registered in a dictionary” (Yasuda; column 2, lines 28-30). The Yasuda system includes a voice recognizer which comprises voice dictionary data, string data, and a dictionary control switch for selecting either the voice or string dictionary data (Yasuda; column 3, lines 25-41). The dictionary files are stored in an external storage device which includes a master dictionary (Yasuda; column 3; lines 42-43). “The master dictionary is generated by extracting a word used for at least two application programs. Therefore, there is no duplicate word included in at least two application dictionaries” (Yasuda; column 3, lines 46-49). When running a new application program, a user transmits new string data of the new application to the voice recognizer, but “does not have to register all the words required for the new application program, and he/she can omit the registering of some words which are included in the master dictionary” (Yasuda; column 3, lines 54-57).

Applicants respectfully submit that claims 6 and 19 are patentable over the Yasuda reference. Specifically, Yasuda fails to teach or suggest, among other things, a “method for automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts,” as required by claims 6 and 19 of the present invention. The term “spoken

language interface" is intended to refer to a collection of operable features that allows a user to interact with the application, as previously explained. Yasuda addresses a completely different problem than that of the present invention, namely, performing standard pattern comparing for eliminating duplicate data entries between two or more application program dictionaries. The application dictionary files associated with the master dictionary in Yasuda are clearly distinguishable from a "spoken language interface," as required by the claimed invention. Moreover, although Yasuda may disclose a feature extractor (11 in FIG. 5) for extracting a feature pattern from a user's voice input, Yasuda does not teach or suggest a mechanism which enables the user to interact with the application. Instead, Yasuda discloses a system wherein the user's voice input, from a microphone, is "analyzed to detect a feature thereof, and compared with the voice and string dictionary data to judge whether or not the voice corresponds to the voice and string dictionary data. The judging result is output as a recognition result" (Yasuda; column 4, lines 2-7). Thus, the system of Yasuda merely performs pattern comparison, as the title suggests.

Applicants further assert that the Yasuda reference fails to teach or suggest an "external network transferring the spoken language interface data set to the device," as set forth in claims 6 and 19. In fact, Yasuda fails to teach or suggest any transfer of data over a network. Moreover, Yasuda does not teach or suggest "discovery of the external network," as required by the claimed invention. Even assuming, *arguendo*, that Yasuda impliedly discloses the transfer of data over a network, the type of data communicated by Yasuda differs significantly from that of the claimed invention. Specifically, Yasuda discloses that "[a] user transmits new string dictionary data of a new application program to the voice recognizer 22 via the communication controller 18" (Yasuda; column 3, lines 52-54) and that "the voice dictionary data corresponding to the string is transmitted to the personal computer 20 via the voice dictionary transference controller 16a of the master dictionary controller 16" (Yasuda; column 3, lines 61-65). However, string dictionary data and/or voice dictionary data cannot reasonably be analogized to a "spoken language interface data set," and thus Yasuda fails to teach or suggest this element of claims 6 and 19.

Inasmuch as the Yasuda reference fails to teach or suggest at least "automatically providing a spoken language interface for a user with respect to at least one external network with which the

Application Serial No. 09/460,913

user interacts" and "the external network transferring the spoken language interface data set to the device," as set forth in claims 6 and 19, Applicants respectfully assert that claims 6 and 19 are patentable over Yasuda. Accordingly, favorable reconsideration and allowance of claims 6 and 19 are respectfully solicited.

Claims 7, 8, 16 and 17 stand rejected under §103 as being unpatentable over the Yasuda reference. Specifically, with regard to claims 7 and 17, the Examiner acknowledges that "Yasuda does not teach that the device is in wireless communications with the external network" (present Office Action; page 3, last paragraph). However, the Examiner takes Official Notice that wireless communication is well-known in the art. With regard to claims 8 and 16, the Examiner acknowledges that "Yasuda does not teach . . . a personal data assistant (PDA) operatively coupled to the spoken language interface," but contends that the personal computer disclosed in Yasuda is analogous to the PDA of the present invention (present Office Action; page 4, paragraph 2). Applicants respectfully disagree with the Examiner's contentions.

Applicants respectfully submit that independent claim 16, which is similar in scope to claim 6, is patentable over the Yasuda reference. Yasuda fails to teach or suggest, among other things, an apparatus "for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts," as required by the claimed invention. As previously stated, the application dictionary files associated with the master dictionary (21) in Yasuda are clearly distinguishable from the "spoken language interfaces" of the claimed invention in that the application dictionary files taught by Yasuda are merely stored words that are used by two or more application programs and do not provide or otherwise modify the operable features which allow a user to interact with an application, in contrast to the present application.

Furthermore, Applicants submit that Yasuda fails to teach or suggest a "portable spoken language interface device is operative to: (i) request a spoken language interface data set from the external network upon discovery of the external network; (ii) receive from the external network the spoken language interface data set; and (iii) load the spoken language interface data set into the data structure of the portable spoken language interface device for use by the user interfacing with the

Application Serial No. 09/460,913

external network,” as expressly set forth in claim 16 of the present application. Even assuming, *arguendo*, that the voice recognizer (22) of Yasuda can be analogized to the portable spoken language interface device of the claimed invention, as the Examiner seems to suggest, Applicants respectfully note that Yasuda fails to teach or suggest a mechanism for requesting, receiving and/or loading a spoken language interface data set in a manner consistent with the present application. As previously explained, the dictionary data which is transferred to the voice recognizer from the personal computer (20) in Yasuda, cannot reasonably be analogized to a spoken language interface data set, which provides the operable features that allow a user to interact with a particular application associated therewith.

For at least the above reasons, Applicants respectfully submit that claim 16 is patentable over the Yasuda reference. Accordingly, favorable reconsideration and allowance of claim 16 is respectfully requested.

With regard to claims 7 and 8, which depend from claim 6, and claim 17, which depends from claim 16, Applicants respectfully submit that these claims are also patentable over the prior art of record by virtue of their dependency from their respective independent claims, which are believed to be patentable for at least the reasons set forth above. Accordingly, favorable reconsideration and allowance of claims 7, 8 and 17 are respectfully solicited.

Claim 13 stands rejected under §103 as being unpatentable over Yasuda in view of Balakrishnan. Specifically, the Examiner acknowledges that “Yasuda does not teach, ‘the portable spoken language interface device prompting the user for information . . . , the device being responsive to the spoken utterance for operatively modifying at least one of a predetermined parameter of the device and an application running on the device,’” but contends that Balakrishnan teaches such features (present Office Action; page 4, paragraph 4). While disagreeing with the Examiner’s contention, Applicants respectfully submit that claim 13, which depends from claim 6, is patentable over the prior art by virtue of its dependency from claim 13, which is believed to be patentable for at least the reasons given above. Accordingly, favorable reconsideration and allowance of claim 13 is respectfully solicited.

Claims 6-8, 13, 16, 17 and 19 stand rejected under §103 as being unpatentable over the Kimura reference. Kimura is directed to a dictionary server, for storing dictionary data, and management system, for updating a dictionary contained in a language processing system (Kimura; column 11, lines 46-47). The dictionary management system communicates between the dictionary server and the language processing system. The dictionary server maintains various dictionary sets that are hierarchically composed (Kimura; column 4, lines 47-51). A user can choose to download desired dictionary data (Kimura; FIG. 4A) from the dictionary server, or “upload the user dictionary data registered in the dictionary use system 103 by the user to the dictionary server 101” (Kimura; column 4, lines 65-67).

With regard to independent claims 6, 16 and 19, which are of similar scope, Applicants respectfully submit that these claims are patentable over the Kimura reference. Specifically, Kimura fails to teach or suggest providing any “spoken language interface,” as required by the claimed invention. As stated above in connection with the Yasuda reference, the spoken language interface of the present application refers to a collection of operable features that allows a user to interact with the application. In contrast, Kimura relates only to dictionary data, which consists of a set of words used by a language processing system. Such data is merely acted upon by an application program and does not provide a user with a means of interacting with the application program itself. Moreover, as previously explained, the spoken language interface of the present application is intended to refer to the general act of speaking to a machine, listening to a machine, and/or interacting with a machine through utterances or audible expressions, and does not refer to a particular lingual type (e.g., Japanese, English, etc.). Thus, the spoken language interface data set of the claimed invention is clearly distinguishable from the dictionary data taught by Kimura.

Kimura also fails to teach or suggest any spoken interaction between the user and the application. Specifically, the Examiner acknowledges that “Kimura does not explicitly teach, that the language is the spoken language” (present Office Action; page 5, paragraph 3). However, the Examiner contends that it would have been obvious to use a spoken language so as to create correct pronunciation of that language (present Office Action; page 5, paragraph 3). Applicants respectfully disagree with the Examiner’s contention and note that it is not the objective of the present invention

Application Serial No. 09/460,913

to correct the pronunciation of a word. As previously explained, the present invention does not relate to language in a lingual sense, but rather provides a spoken language interface for controlling, or otherwise interacting with, the operable features which are built into an application at least in part by user utterances. Inasmuch as Kimura is not directed to at least providing a spoken language interface, which is an essence of the claimed invention, Applicants believe that Kimura is nonanalogous art. For at least the reasons given above, Applicants respectfully submit that independent claims 6, 16 and 19 are patentable over the Kimura reference. Accordingly, favorable reconsideration and allowance of these claims are respectfully requested.

With regard to claims 7, 8 and 13, which depend from claim 6, and claim 17, which depends from claim 16, Applicants respectfully submit that these claims are also patentable over the prior art of record by virtue of their dependency from their respective independent claims, which are believed to be patentable for at least the reasons set forth above. Accordingly, favorable reconsideration and allowance of claims 7, 8, 13 and 17 are respectfully solicited.

In view of the foregoing, Applicants believe that pending claims 1-19, as amended, are in condition for allowance and respectfully request withdrawal of the §112, §102 and §103 rejections.

Attached hereto is a marked-up version of the changes made to the claims by the present Amendment. The attachment is captioned "Version with markings to show changes made."

Respectfully submitted,



Date: March 7, 2002

Wayne L. Ellenbogen  
Attorney for Applicant(s)  
Reg. No. 43,602  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-7662

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS

Claims 1 and 18 have been amended as follows:

1. (Twice amended) In apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, the device having at least one application associated therewith, the spoken language interface apparatus comprising: (A) an audio input system for receiving speech data provided by the user; (B) an audio output system for outputting speech data to the user; (C) a speech decoding engine for generating [a decoded] an output in response to spoken utterances; (D) a speech synthesizing engine for generating a synthesized speech output in response to text data; (E) a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech decoding engine and the speech synthesizing engine; and (F) at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device; wherein: (i) the dialog manager enables connection between the input audio system and the speech decoding engine such that the spoken utterance provided by the user is provided from the input audio system to the speech decoding engine; (ii) the output generated by the speech decoding engine [decodes the spoken utterance to generate a decoded output which] is returned to the dialog manager; (iii) the dialog manager uses the [decoded] output generated by the speech decoding engine to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found; (iv) the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith; (v) the application of the device, on processing that element, provides a reference to an interface element to be spoken; (vi) the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a synthesized output that expresses that element;

and (vii) the audio output system audibly presenting the synthesized output to the user; a method for modifying a data structure containing the at least one user interface data set, comprising:

adding a new application to the device;

generating a second user interface data set in accordance with the new application, the second user interface data set representing spoken language interface elements and data recognizable by the new application;

transferring the second user interface data set from the device to the apparatus; and loading the second user interface data set into the data structure of the apparatus.

18. (Amended) The apparatus of claim 16, wherein the portable spoken language interface device comprises a personal speech assistant (PSA), the PSA comprising:

an audio input system for receiving speech data provided by the user;

an audio output system for outputting speech data to the user;

a speech decoding engine for generating [a decoded] an output in response to spoken utterances;

a speech synthesizing engine for generating a synthesized speech output in response to text data;

a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech decoding engine and the speech synthesizing engine; and

at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device;

wherein:

the dialog manager enables connection between the input audio system and the speech decoding engine such that the spoken utterance provided by the user is provided from the input audio system to the speech decoding engine;

the output generated by the speech decoding engine [decodes the spoken utterance to generate a decoded output which] is returned to the dialog manager;

the dialog manager uses the [decoded] output generated by the speech decoding engine to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found;

the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith;

the application of the device, on processing that element, provides a reference to an interface element to be spoken;

the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a synthesized output that expresses that element; and

the audio output system audibly presents the synthesized output to the user.